

BREVET D'INVENTION

CERTIFICAT D'UTILITÉ - CERTIFICAT D'ADDITION

COPIE OFFICIELLE

Le Directeur général de l'Institut national de la propriété industrielle certifie que le document ci-annexé est la copie certifiée conforme d'une demande de titre de propriété industrielle déposée à l'Institut.

Fait à Paris, le 27 JAN. 2003

DOCUMENT DE PRIORITÉ

PRÉSENTÉ OU TRANSMIS
CONFORMÉMENT À LA
RÈGLE 17.1.a) OU b)

Pour le Directeur général de l'Institut
national de la propriété industrielle
Le Chef du Département des brevets

Martine PLANCHE

INSTITUT
NATIONAL DE
LA PROPRIÉTÉ
INDUSTRIELLE

SIEGE
26 bis, rue de Saint Petersburg
75800 PARIS cedex 08
Téléphone : 33 (0)1 53 04 53 04
Télécopie : 33 (0)1 53 04 45 23
www.inpi.fr

BEST AVAILABLE COPY



26 bis, rue de Saint Pétersbourg
75800 Paris Cedex 08
Téléphone : 01 53 04 53 04 Télécopie : 01 42 94 86 54

BREVET D'INVENTION CERTIFICAT D'UTILITÉ

Code de la propriété intellectuelle - Livre VI



REQUÊTE EN DÉLIVRANCE 1/2

Remplir impérativement la 2ème page.

Cet imprimé est à remplir lisiblement à l'encre noire

DB 540 W / 190500

REMISE EN DÉLIVRANCE DATE 24 JAN 2002 LIEU 75 INPI PARIS N° D'ENREGISTREMENT 0200872 NATIONAL ATTRIBUÉ PAR L'INPI DATE DE DÉPÔT ATTRIBUÉE PAR L'INPI 24 JAN. 2002 Vos références pour ce dossier (facultatif) PIndF1198/13 FR		Réservé à l'INPI NOM ET ADRESSE DU DEMANDEUR OU DU MANDATAIRE À QUI LA CORRESPONDANCE DOIT ÊTRE ADRESSÉE CABINET ORES 6, Avenue de Messine 75008 PARIS FRANCE	
Confirmation d'un dépôt par télécopie <input type="checkbox"/> N° attribué par l'INPI à la télécopie			
2 NATURE DE LA DEMANDE		Cochez l'une des 4 cases suivantes	
Demande de brevet		<input checked="" type="checkbox"/>	
Demande de certificat d'utilité		<input type="checkbox"/>	
Demande divisionnaire		<input type="checkbox"/>	
Demande de brevet initiale N° _____ Date ____/____/____ ou demande de certificat d'utilité initiale N° _____ Date ____/____/____			
Transformation d'une demande de brevet européen Demande de brevet initiale N° _____ Date ____/____/____			
3 TITRE DE L'INVENTION (200 caractères ou espaces maximum) PROCEDE DE SYNCHRONISATION DE DEUX FLUX DE DONNEES NUMERIQUES DE MEME CONTENU.			
4 DÉCLARATION DE PRIORITÉ OU REQUÊTE DU BÉNÉFICE DE LA DATE DE DÉPÔT D'UNE DEMANDE ANTÉRIEURE FRANÇAISE		Pays ou organisation _____ N° _____ Date ____/____/____ Pays ou organisation _____ N° _____ Date ____/____/____ Pays ou organisation _____ N° _____ Date ____/____/____ <input type="checkbox"/> S'il y a d'autres priorités, cochez la case et utilisez l'imprimé «Suite»	
5 DEMANDEUR		<input type="checkbox"/> S'il y a d'autres demandeurs, cochez la case et utilisez l'imprimé «Suite»	
Nom ou dénomination sociale		TELEDIFFUSION DE FRANCE	
Prénoms			
Forme juridique		Société Anonyme	
N° SIREN			
Code APE-NAF			
Adresse	Rue	10, rue d'Oradour-sur-Glane	
	Code postal et ville	75015	PARIS
Pays		FRANCE	
Nationalité		Française	
N° de téléphone (facultatif)			
N° de télécopie (facultatif)			
Adresse électronique (facultatif)			



BREVET D'INVENTION
CERTIFICAT D'UTILITÉ

REQUÊTE EN DÉLIVRANCE 2/2

REMARQUE DES PIÈCES DATE: 24 JAN 2002 LIEU: 75 INPI PARIS N° D'ENREGISTREMENT: 0200872 NATIONAL ATTRIBUÉ PAR L'INPI		Réservé à l'INPI	
Vos références pour ce dossier : (facultatif)		PJndF1198/13 FR	
6 MANDATAIRE			
Nom		JACQUARD	
Prénom		Philippe	
Cabinet ou Société		CABINET ORES	
N° de pouvoir permanent et/ou de lien contractuel			
Adresse	Rue	6, Avenue de Messine	
	Code postal et ville	75008	PARIS
N° de téléphone (facultatif)		01.45.62.75.00 / 01.45.62.69.99	
N° de télécopie (facultatif)		01.45.62.04.86 / 01.45.63. 04.47	
Adresse électronique (facultatif)		ores@cabinet-ores.com	
7 INVENTEUR (S)			
Les inventeurs sont les demandeurs		<input type="checkbox"/> Oui <input checked="" type="checkbox"/> Non Dans ce cas fournir une désignation d'inventeur(s) séparée	
8 RAPPORT DE RECHERCHE		Uniquement pour une demande de brevet (y compris division et transformation)	
Établissement immédiat ou établissement différé		<input checked="" type="checkbox"/> <input type="checkbox"/>	
Paiement échelonné de la redevance		Paiement en deux versements, uniquement pour les personnes physiques <input type="checkbox"/> Oui <input checked="" type="checkbox"/> Non	
9 RÉDUCTION DU TAUX DES REDEVANCES		Uniquement pour les personnes physiques <input type="checkbox"/> Requête pour la première fois pour cette invention (joindre un avis de non-imposition) <input type="checkbox"/> Requête antérieurement à ce dépôt (joindre une copie de la décision d'admission pour cette invention ou indiquer sa référence):	
Si vous avez utilisé l'imprimé «Suite», indiquez le nombre de pages jointes			
10 SIGNATURE DU DEMANDEUR OU DU MANDATAIRE (Nom et qualité du signataire)		VISA DE LA PRÉFECTURE OU DE L'INPI M. BLANCANEUX	
JACQUARD Philippe MANDATAIRE N° 92-4024 CABINET ORES			

La loi n°78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés s'applique aux réponses faites à ce formulaire. Elle garantit un droit d'accès et de rectification pour les données vous concernant auprès de l'INPI.

PROCEDE DE SYNCHRONISATION DE DEUX FLUX DE DONNEES NUMERIQUES DE MEME CONTENU.

L'invention concerne un procédé de synchronisation de deux flux de données numériques de même contenu, et en particulier d'un flux de référence correspondant par exemple à une émission par un système de télétransmission, et d'un flux reçu avec d'éventuelles dégradations, utilisable notamment en vue d'une évaluation de la qualité de la transmission.

Dans le domaine de la diffusion de signaux audiovisuels, l'introduction du numérique offre de nouvelles perspectives et permet de proposer un plus grand nombre de services aux utilisateurs.

Lors des différentes étapes nécessaires à la diffusion, les signaux sont modifiés. En effet, les contraintes techniques imposées, en termes de débit ou de largeur de bande par exemple, impliquent l'apparition de dégradations caractéristiques lors de conditions de transmission difficiles.

Pour pouvoir assurer une qualité de service, il est nécessaire de développer des outils et instruments qui permettent de mesurer la qualité des signaux et d'estimer l'importance des dégradations introduites le cas échéant. De nombreux procédés de mesure ont été développés dans ce but. La majorité de ceux-ci se base sur une comparaison du signal (nommé signal de référence) présent à l'entrée du système à étudier, avec le signal (nommé signal dégradé) obtenu à la sortie du système. Certaines méthodes dites à "référence réduite" utilisent une comparaison entre des nombres calculés sur la référence et sur le signal dégradé au lieu d'utiliser directement les échantillons du signal. Dans les deux cas, il est nécessaire de synchroniser temporellement les signaux pour effectuer l'évaluation de qualité par comparaison. La Figure 1 illustre le principe général utilisé par ces méthodes.

Si la synchronisation des signaux peut être facilement réalisée en simulation ou lorsque le système à étudier est réduit (cas d'un codeur-décodeur ou "codec" par exemple) et non distribué géographiquement, cela est très différent dans le cas d'un système complexe, notamment dans le cas de la supervision d'un réseau de diffusion. Ainsi, cette étape de synchronisation est souvent une étape critique des algorithmes de qualitométrie.

Au delà des applications de mesure de la qualité dans un réseau de diffusion, le procédé présenté dans ce rapport est applicable à chaque fois qu'une synchronisation temporelle entre deux signaux audio et/ou

vidéo doit être effectuée, en particulier dans le contexte d'un système réparti et étendu.

Diverses techniques peuvent être utilisées pour synchroniser temporellement des signaux numériques. L'objectif est de mettre en correspondance une partie du signal dégradé S_D , avec une partie du signal de référence S_R . La Figure 2 illustre cette problématique dans le cas de deux signaux audio. Il s'agit de déterminer un décalage DEC permettant de synchroniser les signaux.

Dans le cas d'un signal audio, la partie (ou élément) à mettre en correspondance, est une fenêtre temporelle, c'est à dire une durée arbitraire T du signal.

On peut répartir les méthodes existantes en trois classes :

- Approche par corrélation dans le domaine temporel : C'est l'approche la plus courante. Elle consiste à comparer les échantillons des deux signaux audio S_R et S_D à synchroniser, en se basant sur leur contenu. Ainsi, l'utilisation de la fonction d'intercorrélation normalisée par exemple, entre S_R et S_D , permet de rechercher le maximum de ressemblance sur une plage temporelle donnée T , par exemple de plus ou moins 60 ms, soit 120 ms. La précision de synchronisation obtenue est potentiellement à l'échantillon près.

- Approche par corrélation dans le domaine temporel utilisant des signaux marqueurs : les méthodes qui utilisent ce principe visent à pallier la nécessité de variations significatives dans le signal. Pour cela, un signal marqueur spécifique est inséré dans le signal audio S_R . Ce signal est conçu pour permettre une synchronisation robuste. Ainsi, une procédure d'intercorrélation identique à la précédente peut être appliquée entre les signaux marqueurs extraits des signaux S_R et S_D à synchroniser, ce qui permet en principe une synchronisation robuste quel que soit le contenu du signal audio.

Afin d'être applicable, l'opération d'insertion du signal marqueur doit être telle que le contenu du signal audio soit modifié de la façon la moins perceptible possible. Plusieurs techniques sont utilisables, dont le "watermarking", pour insérer ces signaux marqueurs ou motifs spécifiques.

- Synchronisation par des repères temporels : Les méthodes de cette classe ne sont utilisables que dans le cas où les signaux sont associés à des repères temporels. Ainsi, le processus repose sur la

recherche, pour chaque repère du signal de référence, celui qui est le plus proche dans la suite des repères temporels associée au signal dégradé.

Un procédé de synchronisation de signaux performant se caractérise par un compromis entre :

- 5 - sa précision, c'est à dire l'erreur maximale commise en synchronisant les deux signaux. En particulier, le procédé peut être sensible au contenu des signaux ,
- sa complexité de calcul,
- et enfin le volume de données nécessaires pour effectuer
- 10 la synchronisation.

Le principal inconvénient des techniques les plus classiquement utilisées (l'approche par corrélation mentionnée ci-dessus) est la puissance de calcul nécessaire. En effet, celle-ci devient en effet très importante dès que la plage de recherche T augmente (Figure 2). De plus,

15 l'autre inconvénient majeur est la nécessité d'un contenu évoluant de manière significative en permanence. Suivant le type de signaux analysés, cette hypothèse n'est pas toujours vraie. Le contenu des signaux a donc une influence directe sur la performance de la méthode. Par ailleurs, pour utiliser ce type d'approche sur les signaux temporels complets, il est nécessaire de

20 disposer des deux signaux S_R et S_D au point de comparaison : c'est une contrainte très forte, qu'il n'est pas possible de satisfaire pour certaines applications comme la surveillance d'un réseau de diffusion opérationnel.

La seconde approche (corrélation avec signaux marqueurs) a pour caractéristique de modifier le contenu du signal audio afin d'y insérer des

25 signaux marqueurs, sans garantie de l'impact de cette opération sur la qualité : le procédé de mesure influe donc sur la mesure elle-même. Quelle que soit la performance pour la synchronisation de deux signaux, cette approche ne convient pas toujours à une application réelle d'évaluation de la qualité.

30 Enfin, dans le cas d'une synchronisation par des repères temporels, c'est la nécessité de disposer des repères temporels qui est l'inconvénient majeur. Seules quelques applications peuvent utiliser une telle technique. En effet, la précision des repères temporels n'est pas toujours satisfaisante.

35 Dans le cadre de la surveillance d'un réseau de diffusion et en raison des contraintes multiples subies par les signaux transportés, et des

multiples équipements traversés (codeurs, multiplexeurs, trans-multiplexeurs, décodeurs, ...), il n'y a pas de relation très stricte entre les signaux audio et les repères temporels. Cette solution n'atteint alors pas la précision nécessaire pour une application de mesure de la qualité avec référence.

5 La présente invention a pour but de définir un procédé permettant de réaliser une synchronisation, d'un niveau de précision qui peut être choisi, d'une complexité plus faible que les méthodes existantes, en combinant les avantages de plusieurs approches. Une synchronisation « grossière » délimite selon l'invention une plage d'erreur, dont la durée est
10 compatible avec l'utilisation ultérieure de méthodes classiques de synchronisation « fine » si une précision extrême est requise.

 L'originalité de la méthode proposée est de réaliser une synchronisation à partir d'au moins un paramètre caractéristique, calculé sur les signaux S_D et S_R , définissant une trajectoire multidimensionnelle, pour en
15 déduire une synchronisation des signaux eux-mêmes. La méthode faisant appel au contenu des signaux, ce contenu temporel doit varier en permanence pour assurer une synchronisation optimale, de même que la corrélation temporelle de l'état de l'art antérieur. L'avantage du procédé est de faire une corrélation grâce une trajectoire multidimensionnelle, obtenue
20 notamment en combinant plusieurs paramètres caractéristiques, qui permet d'obtenir une fiabilité supérieure à celle des procédés connus.

 Un avantage fondamental du procédé proposé par l'invention est de ne nécessiter que peu de données pour réaliser une synchronisation, ce qui est très utile dans le cadre de la supervision d'un réseau de diffusion.
25 En effet, dans ce contexte, il n'est généralement pas possible d'avoir les deux signaux complets S_R et S_D au même endroit. Par conséquent, il n'est pas possible d'utiliser une approche par corrélation temporelle classique. De plus, dans le cadre d'une application de qualimétrie, la seconde approche par corrélation avec signaux marqueurs n'est pas facilement applicable car elle
30 agit sur la qualité des signaux. Au contraire, le procédé de synchronisation selon l'invention est compatible avec des techniques de qualimétrie reposant sur la comparaison de paramètres calculés sur les signaux. Les données représentatives de ce ou ces paramètres caractéristiques sont facilement transportables par une liaison numérique jusqu'aux points de
35 comparaison. Cette liaison numérique utilise avantageusement le même canal de transmission que le signal audio ; alternativement, une liaison numérique

dédiée peut être choisie. Dans un mode particulier de réalisation pour une application de mesure de la qualité, les données utilisées pour réaliser la synchronisation sont ainsi issues d'un ou plusieurs paramètres de mesure de la qualité. Par ailleurs, une synchronisation dite grossière est réalisée à partir de données D1 et D2 calculées à intervalles de $\Delta = 1024$ échantillons audio. Une synchronisation fine peut être réalisée avec des données D1 calculées à intervalles de $\Delta = 1024$ échantillons audio, et avec des données D2 calculées à intervalles de $r < \Delta$, par exemple $r = 32$ échantillons audio. Ainsi, le procédé permet dans ce cas de réaliser une synchronisation fine 32 fois plus précise que l'intervalle de transmission des paramètres de mesure de la qualité.

Le procédé s'intègre ainsi naturellement dans un système de surveillance de la qualité en télévision numérique sur un réseau de diffusion opérationnel. Mais il est applicable à chaque fois qu'une synchronisation temporelle entre deux signaux doit être effectuée.

Le procédé proposé permet ainsi de réaliser la synchronisation avec une précision que l'on peut choisir et ainsi obtenir une plage d'incertitude très réduite. Avantageusement, il utilisera au moins en partie des paramètres déjà calculés pour étudier la qualité du signal. Le fait de pouvoir partir d'une plage de recherche étendue est également un atout, d'autant plus que la robustesse de la synchronisation augmente avec la largeur de cette plage de départ.

Le procédé proposé n'impose donc pas d'avoir de repères temporels externes aux signaux audio. Le signal à synchroniser n'a pas besoin non plus d'être modifié, ce qui est une condition importante dans une application de mesure de la qualité.

L'invention concerne ainsi un procédé de synchronisation entre deux flux de données numériques de même contenu caractérisé en ce qu'il met en œuvre :

a) pour chacun des deux flux de données numériques S_1 et S_2 , générer à des intervalles donnés d'au moins deux nombres caractéristiques exprimant au moins un paramètre caractéristique de leur contenu ;

b) pour chacun des deux flux S_1 et S_2 , générer à partir desdits nombres de points D_1 et D_2 associés respectivement à chacun desdits flux et représentant au moins un dit paramètre caractéristique dans un espace à au moins deux dimensions, les points D_1 d'une part et D_2 d'autre part, qui

appartiennent à une plage temporelle T , définissant des trajectoires représentatives des flux de données S_1 et S_2 à synchroniser ;

c) décaler l'une par rapport à l'autre les plages temporelles de durée T affectées aux flux de données numériques S_1 et S_2 en calculant un critère de superposition desdites trajectoires dont une valeur optimale représente la synchronisation recherchée ;

d) choisir comme valeur représentative de la synchronisation, le décalage entre les plages temporelles correspondant à ladite valeur optimale.

Avantageusement, le procédé est caractérisé en ce qu'un des flux de données numériques est un flux de référence S_1 et en ce que l'autre flux de données est un flux S_2 reçu à travers un système de transmission et en ce que les nombres caractéristiques de flux de référence S_1 sont transmis avec celui-ci, alors que les nombres caractéristiques du flux reçu S_2 sont calculés à la réception.

Selon une première variante, le procédé est caractérisé en ce que c) met en œuvre :

c1) calculer une distance D entre d'une part une première trajectoire représentée par les points D_1 appartenant à une première plage temporelle de durée T , et d'autre part, une deuxième trajectoire représentée par les points D_2 appartenant à une deuxième plage temporelle de durée T , ladite distance D constituant ledit critère de superposition ;

c2) décaler l'un par rapport à l'autre lesdites première et deuxième plages temporelles de durée T jusqu'à obtention d'un minimum de la distance D qui constitue ladite valeur optimale ;

La distance D peut être une moyenne arithmétique des distances d , par exemple des distances euclidiennes, entre les points correspondants D_1 , D_2 des deux trajectoires.

Selon une deuxième variante, le procédé est caractérisé en ce que c) met en œuvre :

c1) calculer une fonction de corrélation entre les points correspondants (D_1 , D_2) des deux trajectoires, ladite fonction de corrélation constituant ledit critère de superposition ;

c2) décaler l'un par rapport à l'autre lesdites première et deuxième plages temporelles de durée T jusqu'à obtention d'un minimum de la fonction de corrélation qui constitue ladite valeur optimale.

Selon une troisième variante, le procédé est caractérisé en ce que c) met en œuvre :

c1) transformer chaque trajectoire en une suite d'angles entre des segments successifs définis par les points de la trajectoire.

5 c2) décaler l'un par rapport à l'autre lesdites première et deuxième plages temporelles de durée T jusqu'à obtenir un minimum entre les différences entre les valeurs d'angles obtenues pour des segments homologues des deux trajectoires, ledit minimum constituant ladite valeur optimale.

10 Le procédé peut être caractérisé en ce que c) met en œuvre :

c1) transformer les deux trajectoires en une suite de surfaces interceptées par les segments successifs définis par les points desdites trajectoires, la surface totale interceptée constituant ledit critère de superposition.

15 c2) décaler l'un par rapport à l'autre les plages temporelles de durée T jusqu'à obtenir un minimum de ladite surface totale interceptée, qui constitue ladite valeur optimale.

Un dit intervalle donné peut être égal à Δ pour un des deux flux de données et à $r < \Delta$ pour l'autre flux de données et ce afin d'améliorer la précision de la synchronisation.

20 Le procédé peut être caractérisé en ce que la génération desdits nombres caractéristiques met en œuvre pour un flux audio de référence et pour le flux audio transmis, les étapes suivantes :

25 a) calculer pour chaque fenêtre temporelle la densité spectrale de puissance du flux audio et lui appliquer un filtre représentatif de l'atténuation de l'oreille interne et moyenne pour obtenir une densité spectrale filtrée,

30 b) calculer à partir de la densité spectrale filtrée les excitations individuelles à l'aide de la fonction d'étalement fréquentiel dans l'échelle basilaire,

c) déterminer à partir desdites excitations individuelles la sonie compressée à l'aide d'une fonction modélisant la sensibilité non linéaire en fréquence de l'oreille, pour obtenir des composantes basilaire,

35 d) séparer les composantes basilaire en n classes (par exemple avec $n \leq 5$), de préférence en trois classes, et calculer pour chaque classe un nombre C représentant la somme des fréquences de cette classe,

les nombres caractéristiques étant constitués par les nombres C. Alternativement, il y a $n' < n$ nombres caractéristiques qui sont générés à partir desdits nombres C. n est choisi nettement inférieur au nombre d'échantillons audio d'une fenêtre temporelle, par exemple inférieur à 0,01 fois ce nombre d'échantillons.

Le procédé peut être caractérisé en ce que la génération d'un dit nombre caractéristique met en œuvre, pour un flux audio de référence et pour un flux audio transmis, les étapes suivantes :

a) calculer N coefficients d'un filtre de prédiction par une modélisation autorégressive,

b) déterminer dans chaque fenêtre temporelle le maximum du résidu par différence entre le signal prédit à l'aide du filtre de prédiction et le signal audio, ledit maximum du résidu de prédiction constituant un dit nombre caractéristique.

Le procédé peut être caractérisé en ce que la génération de dits nombres caractéristiques met en œuvre pour un flux audio de référence et pour un flux audio transmis, les étapes suivantes :

a) calculer pour chaque fenêtre temporelle la densité spectrale de puissance du flux audio et lui appliquer un filtre représentatif de l'atténuation de l'oreille interne et moyenne, pour obtenir une fonction d'étalement fréquentiel dans l'échelle basilaire,

b) calculer des excitations individuelles à partir de la fonction d'étalement fréquentiel dans l'échelle basilaire,

c) obtenir à partir desdites excitations individuelles la sonie compressée à partir d'une fonction modélisant la sensibilité non linéaire en fréquence de l'oreille, pour obtenir des composantes basilaire ,

d) calculer à partir desdites composantes basilaire N' coefficients de prédiction d'un filtre de prédiction par une modélisation autorégressive,

e) générer pour chaque fenêtre temporelle au moins un nombre caractéristique à partir d'au moins un des N' coefficients de prédiction.

Les nombres caractéristiques peuvent être constitués par entre 1 et 10 desdits coefficients de prédiction, et préférentiellement entre 2 et 5 desdits coefficients.

Pour un signal audio numérique, un nombre caractéristique peut être la puissance instantanée et/ou la densité spectrale de puissance et/ou la bande passante.

Pour un signal vidéo numérique, un nombre caractéristique
 5 peut être le coefficient continu de la transformée d'au moins une zone d'une image appartenant au flux de données par une transformée linéaire et orthogonale, par blocs ou globale, et/ou le contraste d'au moins une zone de l'image, et/ou l'activité spatiale SA d'au moins une zone d'une image ou son
 10 activité temporelle (celle-ci étant définie par comparaison avec une image précédente), et/ou la moyenne de la luminosité d'au moins une zone d'une image.

Les points peuvent être générés à partir d'au moins deux nombres caractéristiques issus d'un seul paramètre caractéristique.

Alternativement lesdits points peuvent être générés à partir
 15 d'au moins deux nombres caractéristiques issus d'au moins deux paramètres caractéristiques audio et/ou vidéo.

Le procédé peut être caractérisé en ce que le flux de données comporte des données vidéo et des données audio et en ce qu'il met en œuvre une première synchronisation vidéo à partir de points D'_1 et D'_2
 20 associées à au moins un paramètre caractéristique vidéo correspondant audit flux vidéo, et une deuxième synchronisation audio à partir de points D''_1 et D''_2 associés à au moins un paramètre caractéristique audio correspondant audit flux audio.

Il peut être alors caractérisé en ce qu'il présente une étape de
 25 détermination du décalage de synchronisation entre le flux vidéo et le flux audio par différence entre lesdits décalages obtenus pour le flux vidéo et pour le flux audio.

D'autres caractéristiques et avantages de l'invention apparaîtront mieux à la lecture de la description, en liaison avec les dessins
 30 annexés dans lesquels :

- la figure 1 illustre l'architecture d'un système de mesure de qualité d'un signal audio selon l'Art Antérieur ;
- la figure 2 illustre la problématique de la synchronisation de signaux audio ;

- la figure 3 illustre une augmentation de la précision de la synchronisation susceptible d'être obtenue dans le cadre de la présente invention ;

- la figure 4 donne un exemple de deux trajectoires
5 bidimensionnelles de signaux audio à synchroniser, dans le cas où $r = \Delta / 2$;

- les figures 5 et 6 illustrent deux variantes de synchronisation entre deux trajectoires affectées aux deux flux de données ;

- la figure 7 est un synoptique d'un procédé de synchronisation selon l'invention (synchronisation par trajectoire) ;

10 - les figures 8 à 10 illustrent une synchronisation selon l'invention, le paramètre significatif étant un paramètre perceptuel audio, les figures 10a et 10b illustrant la situation avant et après synchronisation de deux trajectoires.

15 - la figure 11 illustre une mise en œuvre du procédé utilisant comme paramètre caractéristique une modélisation autorégressive du signal, avec coefficients de prédiction linéaire.

La première étape du procédé correspond au calcul d'au moins deux nombres caractéristiques à partir d'un ou de plusieurs paramètres caractéristiques sur toutes les fenêtres temporelles des signaux à
20 synchroniser, sur la plage de synchronisation voulue : chaque nombre est donc calculé toutes les durées Δ (Cf. Figures 2, et 3), ce qui donne $N = T/\Delta$ paramètres. Ce ou ces nombres doivent être, si possible, simples à calculer pour ne pas demander trop de puissance de calcul. Chaque paramètre caractéristique peut être de nature quelconque et peut être représenté par
25 exemple par un seul nombre. Par exemple, pour un signal audio un paramètre caractéristique du contenu du signal peut être la bande passante.

Le fait de disposer des paramètres seulement toutes les durées Δ permet de réduire grandement la quantité de données issues du signal de référence S_R nécessaires à la synchronisation. Cependant, la
30 précision de la synchronisation qui sera réalisée est nécessairement limitée : l'incertitude par rapport à une synchronisation idéale, c'est à dire à l'échantillon du signal près, est de $\pm \Delta / 2$. Dans le cas où cette incertitude est trop grande, une alternative est de diminuer la période Δ : cependant, cette modification est rarement possible car d'une part le calcul du ou des nombres
35 caractéristiques est remis en cause, et d'autre part cela aboutit à augmenter la quantité de données nécessaire à la synchronisation.

Dans le mode particulier de réalisation où les paramètres servent également à évaluer la qualité par comparaison des paramètres P_1 et P'_1 , toute erreur de synchronisation supérieure à la valeur r_0 dite résolution du paramètre ne permettra pas d'estimer les dégradations introduites (Situation A de la Figure 3).

Afin d'obtenir une précision de synchronisation arbitraire, d'une valeur d'incertitude r qui peut être par exemple inférieure à $\Delta/2$, tout en n'augmentant pas la quantité de données extraites du signal de référence, le procédé peut mettre en œuvre un calcul des nombres caractéristiques avec une résolution temporelle plus fine. Pour cela, les paramètres sont calculés toutes les durées $r < \Delta$, sur le second signal à synchroniser dit « dégradé », ce qui correspond à Δ/r paramètres P_1^i pour un paramètre P_1 . La complexité de calcul augmente, en passant de T/Δ à T/r fenêtres de calcul, mais seulement sur le signal reçu. La situation B Figure 3 illustre la méthode utilisée. Par exemple, r est un sous-multiple de Δ .

Notations :

- T : plage de recherche de synchronisation. T est un multiple de Δ .
- r_0 : erreur / incertitude de synchronisation maximale admissible
- e : erreur de synchronisation
- Δ période de calcul des paramètres sur le signal.
- P_k : paramètre calculé sur le premier signal dit de "référence" S_R . k est un indice temporel repérant à quelle période de calcul Δ correspond P_k .
- P'_k : paramètre calculé sur le second signal dit "dégradé" S_D . k est un indice temporel repérant à quelle période de calcul Δ correspond P_k
- P_k^i : paramètre calculé sur le second signal dit "dégradé" S_D . k est un indice temporel repérant à quelle période de calcul Δ correspond P_k . i est un sous-indice temporel repérant un nombre de durées r à l'intérieur de la période p compris entre 1 et Δ/r .

Remarque : Toutes les durées correspondent à un nombre entier d'échantillons du signal audio ou vidéo.

Pour définir une ou plusieurs coordonnées, la deuxième étape applique un traitement sur les paramètres. Un jeu de β coordonnées est

calculé pour chaque jeu de paramètres P_k ou P_k^1 obtenu sur la fenêtre k de durée Δ correspondant par exemple à 1024 échantillons, du signal référence ou dégradé respectivement.

- Le but premier de cette étape est d'obtenir des valeurs de coordonnées pertinentes pour réaliser une synchronisation, et avec des bornes et limites données. Ainsi, chaque coordonnée est obtenue à partir d'une combinaison des nombres caractéristiques disponibles. De plus, cette étape permet de réduire le nombre de dimensions et donc de simplifier les opérations ultérieures.

- 5 Dans un mode de réalisation préféré, deux coordonnées doivent être obtenues ($\beta = 2$). Par exemple, si deux paramètres caractéristiques sont utilisés, chacun d'eux peut servir pour déterminer une coordonnée. Alternativement, un nombre plus élevé de nombres caractéristiques peut être utilisé ; des traitements peuvent être réalisés pour
- 10 passer un nombre moins élevés de nombres, par exemple à deux coordonnées et ceux-ci s'interprètent alors comme une projection d'un espace à autant de dimensions que de nombres caractéristiques vers un espace à par exemple 2 coordonnées.

- La troisième étape correspond à la construction de la
- 20 trajectoire (Figure 4). La trajectoire définit une signature d'un segment du signal audio sur la durée T par une suite de points dans un espace à autant de dimensions que de coordonnées. L'utilisation d'un espace à deux dimensions ou plus permet de construire une trajectoire bien particulière, qui permet d'obtenir pour la synchronisation une fiabilité et une précision élevées.

- 25 Après ces trois étapes, la synchronisation des signaux se résume à la synchronisation de deux trajectoires (ou courbes paramétrées par le temps) dans un espace à deux dimensions ou plus :

- la première trajectoire est définie par les points R_k , obtenus à partir des nombres significatifs P_k calculés toutes les durées Δ sur la plage temporelle T . Il y a $N = T/\Delta$ points R_k .

- la seconde trajectoire est définie par les points $D_k = D_k^1$, obtenus à partir des nombres significatifs $P_k = P_k^1$ calculés toutes les durées Δ sur la plage T . Il y a $N' = N = T/\Delta$ points D_k .

- Lorsque une durée $r < \Delta$ est utilisée pour calculer les
- 35 paramètres P_k^1 , la trajectoire est définie par les points D_k^1 , au nombre de $N' = T/r$ points.

A cet effet, un critère de ressemblance entre deux trajectoires de N points (ou de N et N' points) est mis en oeuvre. Les méthodes suivantes sont données à titre d'exemple :

La première méthode proposée est une minimisation d'une distance entre les deux trajectoires.

L'idée de base est de calculer, sur une portion de la trajectoire, une distance. En fonction de la plage de désynchronisation maximale possible des courbes, correspondant aux signaux audio, ou vidéo, une portion appropriée de chaque trajectoire est sélectionnée.

Sur ces portions, un cumul *Diff* des distances d entre les sommets R_k et $D_{k+\delta}$ ou $D_{k+\delta}^i$ des courbes est calculé respectivement par les relations (1) et (2) ci-après, en appliquant des décalages δ successifs afin de trouver le décalage minimisant la distance *Diff* entre trajectoires.

La Figure 4 illustre le calcul sur un exemple, avec des points définis par deux coordonnées, dans un espace à $\beta = 2$ dimensions. Sur le signal dit « dégradé », les paramètres sont calculés toutes les durées $r=\Delta/2$, c'est à dire avec une résolution deux fois plus fine que sur le premier signal

La distance *Diff* donne l'écart entre les deux trajectoires. La moyenne arithmétique des distances sommet à sommet est un mode de réalisation préféré, mais un autre calcul de distance est également applicable.

$$Diff(\delta) = \alpha \sqrt{\frac{1}{N} \sum_{k=1}^N [d(D_k, R_{k+\delta})]} \quad \text{où } \alpha D = 1.. \infty \text{ et } N = T/\Delta \quad (1)$$

avec $d(A,B)$ distance entre deux points ou sommets. Cette distance $d(A,B)$ peut également être quelconque. Dans un mode particulier de réalisation, la distance Euclidienne est utilisée :

$$d(A,B) = \alpha d \sqrt{\sum_{j=1}^{\beta} (a_j - b_j)^2} \quad \text{où } \alpha d = 1.. \infty \quad (2)$$

où a_j et b_j sont les coordonnées des points A et B
et β désigne le nombre de coordonnées de chaque point.

Le décalage δ donnant la distance *Diff* minimale correspond à la désynchronisation des courbes, et par conséquent des signaux de départ. Dans l'exemple proposé (Figure 4), le décalage trouvé

sera de 2, soit 2 fois la période Δ de calcul du paramètre initial. La plage de synchronisation sera donc comprise entre :

$$t + 2 * \Delta - \frac{\Delta}{2} \quad \text{et} \quad t + 2 * \Delta + \frac{\Delta}{2} \quad (3)$$

Le deuxième critère proposé est une maximisation d'une
5 corrélation entre les deux trajectoires

Ce critère fonctionne d'une manière similaire au précédent, à la différence qu'il consiste à maximiser la valeur *Correl*. Les relations (1) et (2) sont remplacées par les deux suivantes.

$$Correl(delta) = \sum_{k=1}^N D_k * R_{k+delta} \quad (4)$$

10 où l'opération * est le produit scalaire défini par :

$$A * B = \frac{\sum_{k=1}^N a_j * b_j}{\sqrt{\sum_{k=1}^N a_j^2} \cdot \sqrt{\sum_{k=1}^N b_j^2}} \quad (5)$$

où a_j et b_j sont les coordonnées des points A et B.

15 Les méthodes ci-après conviennent particulièrement au cas où $\beta = 2$ coordonnées

D'autres techniques permettent de rendre le procédé plus robuste à la présence de différences significatives entre les signaux à synchroniser, dues par exemple à des dégradations lors d'une diffusion, à
20 savoir.

- distance entre les angles successifs des trajectoires.

Cette méthode consiste à transformer la trajectoire bidimensionnelle en une suite d'angles mesurés entre des segments successifs définis par les points de la trajectoire. La Figure 5 illustre la
25 définition des angles $\Delta\varphi$.

Le critère utilisé pour synchroniser les deux trajectoires est la minimisation de la relation suivante :

$$Diff(delta) = \sum_{k=1}^{N-1} |\varphi_k - \varphi_{k+delta}| \quad (6)$$

- surface interceptée entre les deux courbes.

5 Cette méthode consiste à transformer la trajectoire bidimensionnelle en une suite des surfaces interceptées par les segments successifs définis par les points de la trajectoire. La Figure 6 illustre la définition des surfaces interceptées S .

Le critère utilisé pour synchroniser les deux trajectoires est la minimisation de la relation suivante :

$$10 \quad S_{\text{Totale}} = \text{somme } S \quad Diff(delta) = \sum_{k=1}^{N-1} |S_{k,k+delta}| \quad (7)$$

15 ◦ Enfin, l'utilisation simultanée de plusieurs critères est possible. Une fois que la valeur $delta$ du décalage de désynchronisation entre les deux signaux a été déterminée par une des méthodes précédentes, les deux signaux peuvent être resynchronisés en appliquant le décalage $delta$ à l'un des signaux. La synchronisation est réalisée, avec une précision déterminée par la cadence de calcul des nombres caractéristiques.

La synoptique d'un processus de synchronisation est donnée à la Figure 7.

20 Lorsque la précision voulue n'est pas atteinte, c'est à dire que la synchronisation est trop « grossière » pour l'application visée, une dernière étape peut consister à affiner le résultat précédent.

25 Une procédure de l'art antérieur peut être appliquée sur la plage d'incertitude de synchronisation Δ ou r , qui est maintenant suffisamment réduite pour être d'une complexité acceptable. Une approche par corrélation dans le domaine temporel, de préférence avec signaux marqueurs, peut être utilisée par exemple.

30 Toutefois, cette étape ne devra être mise à profit que dans certains cas particuliers. En effet, dans le type d'application visée de mesure de qualité, un affinage de la synchronisation n'est en général pas nécessaire car la précision obtenue est suffisante. De plus, les techniques de l'art antérieur nécessitent, ainsi qu'il a été expliqué ci-dessus, de disposer de

données sur les signaux qui ne sont pas facilement transportables dans un système complexe et réparti.

Un des modes particuliers de réalisation correspond à la mise en œuvre de l'invention dans le cadre d'une application de surveillance de la qualité audio dans un réseau de diffusion de télévision numérique. Dans ce cadre, un intérêt majeur de l'invention est d'utiliser les données servant à l'évaluation de qualité pour réaliser la synchronisation, puisque cela évite ou minimise la nécessité de transmettre des données spécifiques à la synchronisation.

Divers nombres caractéristiques (référence réduite), destinés à l'estimation de l'importance des dégradations introduites lors de la diffusion des signaux, sont calculés sur le signal de référence en entrée du réseau. Ces nombres de référence P_R sont transmis via une voie de donnée jusqu'au point de mesure de la qualité. Des nombres caractéristiques P_M sont calculés sur le signal dégradé au niveau de ce point de mesure. L'estimation de la qualité va ainsi se faire en comparant les paramètres P_R et P_M . Pour cela, ils doivent être synchronisés, à partir du ou des paramètres caractéristiques utilisés pour la référence.

L'estimation de la qualité va ainsi se faire en comparant les paramètres P_R et P_M . Pour cela, ils doivent être synchronisés.

Le principe des mesures perceptuelles objectives repose sur la transformation de la représentation physique (pression acoustique, niveau, temps et fréquence) en la représentation psychoacoustique (force sonore, niveau de masquage, temps et bandes critiques ou barks) de deux signaux (le signal de référence et le signal à évaluer) afin de les comparer. Cette transformation s'opère grâce à une modélisation de l'appareil auditif humain (généralement, cette modélisation consiste en une analyse spectrale dans le domaine des Barks suivie des phénomènes d'étalement).

L'exemple ci-après de mise en œuvre du procédé selon l'invention met en œuvre un paramètre caractéristique perceptuel dénommé « Ecart de Comptes Perceptuels ». L'idée originale pour ce paramètre est d'établir une mesure d'uniformité d'une fenêtre du signal audio. Ainsi, un signal sonore dont les composantes fréquentielles sont stables est considéré comme uniforme. Inversement, un bruit « parfait » correspond à un signal qui couvre uniformément toutes les bandes de fréquences (spectre plat). Ce type de paramètre permet donc de caractériser le contenu du signal. Cette

capacité est renforcée par son caractère perceptuel, c'est à dire la prise en compte de caractéristiques du système auditif humain connues par la psychoacoustique.

Plusieurs étapes sont mises en oeuvre pour tenir compte de la psychoacoustique. Elles sont appliquées sur le signal de référence et sur le signal dégradé. Ces étapes sont les suivantes :

- Fenêtrage du signal temporel en blocs, puis, pour chaque bloc, calcul de l'excitation induite par le signal en utilisant un modèle d'audition. Cette représentation des signaux tient compte des phénomènes de la psychoacoustique, et fournit un histogramme dont les comptes sont les valeurs des composantes basilaire. Cela permet de ne prendre en considération que les composantes audibles du signal et donc de se limiter à l'information utile. Pour obtenir cette excitation, les modélisations classiques peuvent être utilisées : atténuation de l'oreille externe et moyenne, intégration selon les bandes critiques et masquages fréquentiels. Les fenêtres temporelles choisies sont d'environ 42 ms (2048 points à 48 kHz) avec un recouvrement de 50%. Cela permet d'obtenir une résolution temporelle de l'ordre de 21 ms.

Plusieurs étapes interviennent pour cette modélisation. Pour la première étape, le filtre d'atténuation de l'oreille externe et moyenne est appliqué à la densité spectrale de puissance, obtenue à partir du spectre du signal. Ce filtre prend également en compte le seuil absolu d'audition. La notion de bandes critiques est modélisée par une transformation de l'échelle fréquentielle en une échelle basilaire. L'étape suivante correspond au calcul des excitations individuelles pour tenir compte des phénomènes de masquage, grâce à la fonction d'étalement dans l'échelle basilaire et à une addition non linéaire. La dernière étape permet d'obtenir la sonie compressée, par une fonction puissance, pour modéliser la sensibilité non linéaire en fréquence de l'oreille, par un histogramme comportant les 109 composants basilaire.

- Les comptes de l'histogramme obtenus sont ensuite périodiquement vectorisés en trois classes pour obtenir une représentation selon une trajectoire. C'est cette trajectoire qui permet de visualiser l'évolution de la structure des signaux et sert à la synchronisation. Cela permet également d'obtenir une caractérisation simple et concise du signal et donc de disposer d'un paramètre de référence (ou paramètre caractéristique).

Plusieurs stratégies existent pour fixer les bornes de ces trois classes : La plus simple est de séparer l'histogramme en trois zones de tailles égales. Ainsi, les 109 composantes basilaies, qui représentent 24 Barks, peuvent être séparées aux indices suivants :

$$5 \quad IS_1 = 36 \text{ soit } z = \frac{24}{109} * 36 = 7,927 \text{ Barks} \quad (8)$$

$$IS_2 = 73 \text{ soit } z = \frac{24}{109} * 73 = 16,073 \text{ Barks} \quad (9)$$

La deuxième stratégie prend en compte les zones de mise à l'échelle (« Scaling ») de BEERENDS. Cela correspond à une compensation du gain entre l'excitation du signal de référence et celle du signal à tester en considérant trois zones sur lesquelles l'oreille réaliserait cette même opération. Ainsi, les bornes fixées sont les suivantes :

$$10 \quad IS_1 = 9 \text{ soit } z = \frac{24}{109} * 9 = 1,982 \text{ Barks} \quad (10)$$

$$IS_2 = 100 \text{ soit } z = \frac{24}{109} * 100 = 22,018 \text{ Barks} \quad (11)$$

La trajectoire est ensuite représentée dans un triangle, nommé triangle des fréquences. Pour chaque bloc on obtient trois comptes C_1 , C_2 et C_3 , donc deux coordonnées cartésiennes selon les formules suivantes :

$$X = C_1/N + \frac{C_2/N}{2} \quad (12)$$

$$Y = C_2/N * \sin(\pi/3) \quad (13)$$

20 avec C_1 : somme des excitations pour les hautes fréquences (au-dessus de S_2)

C_2 : compte associé aux fréquences moyennes (composantes entre S_1 et S_2)

25 et $N = C_1 + C_2 + C_3$: Somme totale des valeurs des composantes.

Un point (X, Y) est donc obtenu pour chaque fenêtre temporelle du signal. Chacune des coordonnées X et Y constitue un nombre caractéristique. Alternativement, C_1 , C_2 et C_3 peuvent être pris comme nombres caractéristiques.

30 Pour une séquence complète, la représentation associée est ainsi une trajectoire paramétrée par le temps, comme le montre la figure 8.

Parmi différentes méthodes possibles pour effectuer la synchronisation des trajectoires, la technique choisie à titre d'exemple est celle basée sur minimisation de la distance entre les points des trajectoires.

Il est important de noter que le calcul du paramètre utilisé dans ce cas pour la synchronisation reste complexe, mais ce paramètre peut être également utilisé pour estimer la qualité du signal. Il doit donc de toute façon être calculé, et ce n'est donc pas une charge supplémentaire ajoutée au calcul lors de la comparaison, d'autant que le calcul relatif à ce paramètre n'est effectué localement que pour le flux numérique reçu.

La figure 9 résume le procédé utilisé pour synchroniser les signaux dans le cadre de la supervision de la qualité de signaux diffusés, avec le paramètre caractéristique ci-dessus.

L'exemple suivant illustre le cas d'un fichier référence (R1), codé puis décodé MPEG2 Layer2 à 128 kbits/s pour obtenir un fichier dégradé (R2). La désynchronisation introduite est de 6000 échantillons. Le décalage trouvé est de 6 fenêtres soit $6 * 1024 = 6144$ échantillons. L'erreur (144) est bien inférieure à la période (1024) du paramètre caractéristique utilisé. Les figures 10a et 10b présentent les trajectoires avant puis après synchronisation :

Avant synchronisation (figure 10a), il n'y a pas de correspondance point à point entre les deux trajectoires. Après synchronisation (figure 10b), la correspondance entre les deux trajectoires est optimale au sens du critère de distance (Cf. relation (1)).

Il n'est, en général, pas nécessaire de faire une synchronisation plus fine. C'est notamment le cas si l'incertitude donnée par la procédure exposée ici est inférieure à l'erreur de synchronisation maximale admissible par le paramètre de mesure de qualité. Pour les paramètres de qualitométrie les plus exigeants, la résolution nécessaire r_0 est de l'ordre de 32 échantillons.

Dans le cas de la figure 10a, la plage initiale est de l'ordre de 120 ms, soit 5760 échantillons à 48 kHz. En utilisant uniquement les nombres caractéristiques disponibles pour l'évaluation de la qualité (tous les 1024 échantillons, soit la durée Δ), une première synchronisation est réalisée avec une incertitude de 1024 échantillons, soit un facteur supérieur à 5 par rapport à 5760, pour une puissance de calcul consacrée à la synchronisation très limitée.

Cependant, par exemple dans une seconde étape, le calcul plus fréquent des paramètres de qualité sur le second signal (dégradé) ($r < \Delta$) permet si on le désire de réduire encore l'erreur de synchronisation à r échantillons.

- 5 Un autre paramètre caractéristique met en œuvre une modélisation autorégressive du signal.

Le principe général de la prédiction linéaire est de modéliser le signal comme étant une combinaison de ses valeurs passées. L'idée est de calculer les N coefficients d'un filtre de prédiction par une modélisation
10 autorégressive (tout pôle). Avec ce filtre adaptatif, il est possible d'obtenir un signal prédit à partir du signal réel. Les erreurs de prédiction ou résidus sont calculés par différence entre ces deux signaux. La présence et la quantité de bruit dans un signal peuvent être déterminées par l'analyse de ces résidus.

La comparaison des résidus obtenus sur le signal de
15 référence et ceux calculés à partir du signal dégradé, et donc des niveaux de bruit, permet d'estimer l'importance des modifications et défauts insérés.

La référence à transmettre correspond au maximum des résidus sur une fenêtre temporelle de taille donnée. Il n'est en effet pas intéressant de transmettre tous les résidus si le débit de la référence veut être
20 réduit.

Pour adapter les coefficients du filtre de prédiction, deux méthodes sont données ci-après à titre d'exemple :

- L'algorithme de LEVINSON-DURBIN qui est décrit par exemple dans l'ouvrage de M. BELLANGER - Traitement numérique du signal
25 - Théorie et pratique (MASSON éd. 1987) p. 393 à 395. Pour l'utiliser, il faut disposer d'une estimation de l'autocorrélation du signal sur un ensemble de N_0 échantillons. Cette autocorrélation est utilisée pour résoudre le système d'équations de Yule-Walker et ainsi obtenir les coefficients du filtre prédicteur. Seules les N premières valeurs de la fonction d'autocorrélation peuvent être
30 utilisées, où N désigne l'ordre de l'algorithme, c'est-à-dire le nombre de coefficients du filtre. Sur une fenêtre de 1024 échantillons, on garde le maximum de l'erreur de prédiction.

- L'algorithme du gradient qui est décrit par exemple dans l'ouvrage précité de M. BELLANGER p. 371 et suivantes. Le principal
35 inconvénient du paramètre précédent est la nécessité, dans le cas d'une implantation sur DSP, de stocker les N_0 échantillons pour estimer

l'autocorrélation, avoir les coefficients du filtre puis calculer les résidus. Ce second paramètre permet d'éviter cela en utilisant un autre algorithme permettant de calculer les coefficients du filtre : l'algorithme du gradient. Celui-ci utilise l'erreur commise pour mettre à jour les coefficients. Les coefficients
 5 du filtre sont modifiés dans la direction du gradient de l'erreur quadratique instantanée, avec le signal opposé.

Une fois les résidus obtenus par différence entre le signal prédit et le signal réel, seul le maximum de leurs valeurs absolues, sur une fenêtre temporelle de taille donnée T , est conservé. Le vecteur référence à
 10 transmettre peut ainsi être réduite à un seul nombre.

Après transmission puis synchronisation, la comparaison consiste en un simple calcul de distance entre les maxima de la référence et du signal dégradé par exemple par différence.

Le principal avantage des deux paramètres est le débit
 15 nécessaire au transfert de la référence. Celui-ci permet de réduire la référence à une valeur pour 1024 échantillons de signal.

Par contre, aucun modèle de la psychoacoustique n'est pris en compte.

Un autre paramètre caractéristique met en œuvre une
 20 modélisation autorégressive de l'excitation basilaire.

Par rapport à la prédiction linéaire classique, cette méthode permet de prendre en compte les phénomènes de la psychoacoustique, afin d'obtenir une évaluation de la qualité perçue. Pour cela, le calcul du paramètre passe par une modélisation de divers principes de l'audition. Une
 25 prédiction linéaire modélise le signal comme étant une combinaison de ses valeurs passées. L'analyse des résidus (ou erreurs de prédiction) permet de déterminer et d'estimer la présence de bruit dans un signal. L'inconvénient majeur lors de l'utilisation de ces techniques est le fait qu'il n'y ait aucune prise en compte des principes de la psychoacoustique. Ainsi, il n'est pas possible
 30 d'estimer la quantité de bruit réellement perçue.

Le procédé reprend le principe général de la prédiction linéaire classique. Elle y intègre en plus les phénomènes de la psychoacoustique pour l'adapter à la sensibilité non linéaire en fréquence (sonie) et en intensité (tonie) de l'oreille humaine.

35 On modifie le spectre du signal, par l'intermédiaire d'un modèle d'audition, avant de calculer les coefficients de la prédiction linéaire

par une modélisation autorégressive (tout pôle). Les coefficients ainsi obtenus permettent de modéliser le signal de façon simple tout en tenant compte de la psychoacoustique. Ce sont ces coefficients de prédiction qui seront transmis et serviront de référence lors de la comparaison avec le signal dégradé.

5 La première partie du calcul de ce paramètre correspond à la modélisation des principes de la psychoacoustique en utilisant les modèles d'audition classiques. La deuxième partie est le calcul des coefficients de prédiction linéaire. La dernière partie correspond à la comparaison des coefficients de prédiction calculés pour le signal de référence et ceux obtenus
10 pour le signal dégradé. Les différentes étapes de cette méthode sont donc les suivantes :

- Fenêtrage temporel du signal puis calcul d'une représentation interne du signal par modélisation des phénomènes de la psychoacoustique. Cette étape correspond au calcul de la sonie compressée,
15 qui est en fait l'excitation induite par le signal au niveau de l'oreille interne. Cette représentation des signaux tient compte des phénomènes de la psychoacoustique, et est obtenue à partir du spectre du signal, en utilisant les modélisations classiques : atténuation de l'oreille externe et moyenne, intégration selon les bandes critiques et masquages fréquentiels. Cette étape
20 du calcul est identique au paramètre décrit précédemment.

- Modélisation autorégressive de cette sonie compressée afin d'obtenir les coefficients d'un filtre RIF de prédiction, tout comme dans une prédiction linéaire classique. La méthode utilisée est celle de l'autocorrélation, par résolution des équations de Yule-Walker. La première étape pour
25 l'obtention des coefficients de prédiction est donc le calcul de l'autocorrélation du signal.

En considérant la sonie compressée comme une puissance spectrale filtrée, il est possible de calculer l'autocorrélation du signal perçu par transformation de Fourier inverse.

30 Une des méthodes pour résoudre ce système d'équations de Yule-Walker et ainsi obtenir les coefficients d'un filtre prédictif est l'utilisation de l'algorithme de Levinson-Durbin.

Ce sont les coefficients de prédiction qui constituent le vecteur de référence à transmettre jusqu'au point de comparaison. Les
35 transformations utilisées lors du calcul final sur le signal dégradé, sont les mêmes que pour la phase initiale sur le signal de référence.

- Estimation des dégradations par le calcul d'une distance entre les vecteurs issus de la référence et du signal dégradé. C'est une comparaison des vecteurs de coefficients définissant des points obtenus pour la référence et pour le signal audio transmis, qui permet d'estimer les dégradations introduites lors de la transmission. Celle-ci doit se faire sur un nombre adapté de coefficients. Plus le nombre est important, plus les calculs peuvent être précis, mais plus le débit nécessaire à la transmission de la référence est élevé. Plusieurs distances peuvent être utilisées pour comparer les vecteurs de coefficients. L'importance relative des coefficients peut par exemple être prise en compte.

Le principe de la méthode se résume selon le schéma suivant (Figure 11).

La modélisation des phénomènes de la psychoacoustique permet d'obtenir 24 composantes basilaies. L'ordre N du filtre de prédiction est de 32. A partir de celles-ci, 32 coefficients de l'autocorrélation sont estimés, ce qui donne 32 coefficients de prédiction dont on ne conserve que 5 à 10 coefficients en tant que vecteur indicateur de qualité, par exemple les 5 à 10 premiers coefficients.

Le principal avantage du paramètre provient de la prise en compte des phénomènes de la psychoacoustique. Pour faire cela, il a été nécessaire d'augmenter le débit nécessaire au transfert de la référence à 5 ou 10 valeurs pour 1024 échantillons de signal (21 ms.), soit un débit de 7,5 à 15 kbits/s.

Le paramètre caractéristique P peut être d'une manière générale toute grandeur obtenue à partir du contenu des signaux numériques et par exemple en vidéo :

- la luminosité de l'image ou d'une zone de celle-ci tel que donnée par le coefficient continu $F(0,0)$ de la transformée en cosinus discrète TCD de l'image, ou de tout autre transformée par blocs, linéaire et orthogonale, par blocs ou globale,

- et/ou le contraste de l'image ou d'une zone de celle-ci, donné par exemple par l'application d'un filtre de Sobel,

- et/ou l'activité de l'image SA de l'image telle que définie par exemple dans la Demande PCT WO 99/18736 de la Demanderesse, et obtenue par une transformation par blocs linéaire et orthogonale (TCD,

transformée de FOURIER, de HAAR, d'HADAMARD, de SLANT, transformée en ondelettes etc...),

- la moyenne de l'image,

et en audio :

5

- la puissance,

- et/ou bien la densité spectrale de puissance telle que définie

dans la demande de Brevet Français FR 2 769 777 déposée le 13 Octobre 1997, et/ou un des paramètres décrits ci-dessus.

10 On remarquera que le paramètre P est susceptible d'être soumis à des dégradations dues à la transmission, mais on observe en pratique, qu'aux niveaux de dégradation généralement observés dans les réseaux de transmission, la synchronisation peut être effectuée à l'aide du procédé selon l'invention.

15 D'une manière générale, une fois la synchronisation acquise, le procédé permet de vérifier qu'elle est bien conservée, de manière à pouvoir remédier en tant que de besoin aux perturbations telles qu'interruption de flux binaire, changement de flux binaire, changement de décodeur, etc..., en resynchronisant les deux signaux numériques E et S.

20 Le procédé décrit est applicable à chaque fois qu'il est nécessaire de synchroniser deux flux numériques. La méthode permet d'obtenir une première plage de synchronisation de largeur suffisamment réduite pour autoriser l'utilisation de méthodes de synchronisation fine classiques, et cela en temps réel.

25 Le procédé exploite avantageusement un ou plusieurs paramètres caractéristiques des signaux à synchroniser qui sont représentés par au moins deux nombres caractéristiques, au lieu de la totalité des signaux. Selon une réalisation préférée, l'utilisation combinée de plusieurs paramètres permet d'assurer une fiabilité de synchronisation supérieure aux techniques antérieures. De plus, l'invention permet de réaliser une
30 synchronisation d'un niveau de précision qui peut être choisi, et d'une complexité plus faible que les méthodes existantes. Cette synchronisation délimite une plage d'erreur, d'une durée autorisant par la suite l'utilisation de méthodes classiques de synchronisation « fine » si une précision plus élevée est requise.

35 L'équipement de mesure implantant le procédé de l'invention, trouve son application notamment dans le domaine des réseaux de diffusion

des signaux audiovisuels numériques afin de permettre une surveillance de la qualité des signaux délivrés par le réseau.

L'invention permet également de réaliser, pour un flux de données incorporant des données audio et vidéo, une synchronisation pour le son et pour l'image. A cet effet, une synchronisation est effectuée pour la vidéo, en calculant un décalage de synchronisation pour la vidéo, et une synchronisation est effectuée pour l'audio, en calculant un décalage de synchronisation pour l'audio. De plus, en comparant la valeur des deux décalages, on peut savoir si lors par exemple d'une transmission, il s'est produit un décalage entre le son et l'image.

REVENDEICATIONS

1. Procédé de synchronisation entre deux flux de données numériques de même contenu caractérisé en ce qu'il met en œuvre :

- a) pour chacun des deux flux de données numériques S_1 et S_2 , une génération à des intervalles donnés d'au moins deux nombres caractéristiques exprimant au moins un paramètre caractéristique de leur contenu ;
- b) pour chacun des deux flux S_1 et S_2 , la génération, à partir desdits nombres, de points D_1 et D_2 représentant au moins un dit paramètre caractéristique dans un espace à au moins deux dimensions, les points D_1 correspondant au flux S_1 d'une part et D_2 correspondant au flux S_2 d'autre part qui sont situés dans une plage temporelle T définissant des trajectoires représentatives des flux de données S_1 et S_2 à synchroniser ;
- c) décaler l'une par rapport à l'autre les plages temporelles de durée T affectées aux flux de données numériques S_1 et S_2 en calculant un critère de superposition desdites trajectoires dont une valeur optimale représente la synchronisation recherchée ;
- d) choisir comme valeur représentative de la synchronisation, le décalage entre les plages temporelles correspondant à ladite valeur optimale.

2. Procédé selon la revendication 1, caractérisé en ce qu'un des flux de données numériques est un flux de référence S_1 et en ce que l'autre flux de données est un flux S_2 reçu à travers un système de transmission et en ce que les nombres caractéristiques du flux de référence S_1 sont transmis avec celui-ci, alors que les nombres caractéristiques du flux reçu S_2 sont calculés à la réception.

3. Procédé selon une des revendications 1 ou 2, caractérisé en ce que c) met en œuvre :

- c1) calculer une distance D entre d'une part une première trajectoire représentée par les points D_1 appartenant à une première plage temporelle de durée T , et d'autre part, une deuxième trajectoire représentée par les points D_2 appartenant à une deuxième plage temporelle de durée T , ladite distance D constituant ledit critère de superposition ;
- c2) décaler l'un par rapport à l'autre lesdites première et deuxième plages temporelles de durée T jusqu'à obtention d'un minimum de la distance D qui constitue ladite valeur optimale ;

4. Procédé selon la revendication 3, caractérisé en ce que la distance D est une moyenne arithmétique des distances d entre les points correspondants D_1 , D_2 des deux trajectoires.

5. Procédé selon la revendication 4, caractérisé en ce que ladite distance d entre les points est une distance euclidienne.

6. Procédé selon une des revendications 1 ou 2, caractérisé en ce que c) met en œuvre :

c1) calculer une fonction de corrélation entre les points correspondants D_1 , D_2 des deux trajectoires, ladite fonction de corrélation constituant ledit critère de superposition ;

c2) décaler l'un par rapport à l'autre lesdites première et deuxième plages temporelles de durée T jusqu'à obtention d'un minimum de la fonction de corrélation qui constitue ladite valeur optimale.

7. Procédé selon une des revendications 1 ou 2, caractérisé en ce que c) met en œuvre :

c1) transformer chaque trajectoire en une suite d'angles entre des segments successifs définis par les points de la trajectoire.

c2) décaler l'un par rapport à l'autre lesdites première et deuxième plages temporelles de durée T jusqu'à obtenir un minimum entre les différences entre les valeurs d'angles obtenues pour des segments homologues des deux trajectoires, ledit minimum constituant ladite valeur optimale.

8. Procédé selon une des revendications 1 ou 2, caractérisé en ce que c) met en œuvre :

c1) transformer les deux trajectoires en une suite de surfaces interceptées par les segments successifs définis par les points desdites trajectoires, la surface totale interceptée constituant ledit critère de superposition.

c2) décaler l'un par rapport à l'autre les plages temporelles de durée T jusqu'à obtenir un minimum de ladite surface totale interceptée, qui constitue ladite valeur optimale.

9. Procédé selon une des revendications précédentes, caractérisé en ce qu'un dit intervalle donné est égal à Δ pour un des flux de données S_1 et à $r < \Delta$ pour l'autre flux de données S_2 .

10. Procédé selon une des revendications précédentes, caractérisé en ce qu'il met en œuvre une première synchronisation en

choisissant un même intervalle donné Δ pour les deux flux de données, puis une deuxième synchronisation pour laquelle l'intervalle donné est égal à Δ pour un des flux de données S1 et à $r < \Delta$ pour l'autre flux de données S2.

11. Procédé selon une des revendications précédentes, caractérisé en ce que la génération desdits nombres caractéristiques met en œuvre pour un flux de données audio de référence et pour un flux de données audio transmis, les étapes suivantes :

a) calculer pour chaque fenêtre temporelle la densité spectrale de puissance du flux audio et lui appliquer un filtre représentatif de l'atténuation de l'oreille interne et moyenne pour obtenir une densité spectrale filtrée,

b) calculer à partir de la densité spectrale filtrée les excitations individuelles à l'aide de la fonction d'étalement fréquentiel dans l'échelle basilaire,

c) déterminer à partir desdites excitations individuelles la sonie compressée à l'aide d'une fonction modélisant la sensibilité non linéaire en fréquence de l'oreille, pour obtenir des composantes basilaire,

d) séparer les composantes basilaire en n classes (avec n nettement inférieur au nombre d'échantillons audio d'une fenêtre temporelle), de préférence en trois classes, et calculer pour chaque classe un nombre C représentant la somme des fréquences de cette classe, ces nombres caractéristiques étant au nombre de n et constitués par un dit nombre C ou étant au nombre de $n' < n$ et étant générés à partir desdits nombres C .

12. Procédé selon une des revendications précédentes, caractérisé en ce que la génération d'un dit nombre caractéristique met en œuvre, pour un flux de données audio de référence et pour un flux de données audio transmis, les étapes suivantes :

a) calculer N coefficients d'un filtre de prédiction par une modélisation autorégressive,

b) déterminer dans chaque fenêtre temporelle le maximum du résidu par différence entre le signal prédit à l'aide du filtre de prédiction et le signal audio, ledit maximum du résidu de prédiction constituant un dit nombre caractéristique.

13. Procédé selon une des revendications précédentes, caractérisé en ce que la génération de dits nombres caractéristiques met en

œuvre pour un flux de données audio de référence et pour un flux de données audio transmis, les étapes suivantes :

- a) calculer pour chaque fenêtre temporelle la densité spectrale de puissance du flux audio et lui appliquer un filtre représentatif de l'atténuation de l'oreille interne et moyenne, pour obtenir une fonction d'étalement fréquentiel dans l'échelle basilaire,
 - b) calculer des excitations individuelles à partir de la fonction d'étalement fréquentiel dans l'échelle basilaire,
 - c) obtenir à partir desdites excitations individuelles la sonie compressée à partir d'une fonction modélisant la sensibilité non linéaire en fréquence de l'oreille, pour obtenir des composantes basilaire ,
 - d) calculer à partir desdits composantes basilaire N' coefficients de prédiction d'un filtre de prédiction par une modélisation auto-régressive,
 - e) générer pour chaque fenêtre temporelle au moins un nombre caractéristique à partir d'au moins des N' coefficients de prédiction.
14. Procédé selon la revendication 13, caractérisé en ce que les nombres caractéristiques sont constitués par, entre 1 et 10 desdits coefficients de prédiction, et préférentiellement entre 2 et 5 desdits coefficients.
15. Procédé selon une des revendications précédentes, caractérisé en ce qu'un nombre caractéristique est la puissance instantanée d'un signal audio.
16. Procédé selon une des revendications précédentes, caractérisé en ce qu'un nombre caractéristique est le densité spectrale de puissance d'un signal audio.
17. Procédé selon une des revendications précédentes, caractérisé en ce qu'un nombre caractéristique est la bande passante d'un signal audio.
18. Procédé selon une des revendications précédentes, caractérisé en ce qu'un nombre caractéristique est le coefficient continu de la transformée par une transformée linéaire et orthogonale d'au moins une partie d'une image appartenant au flux de données, cette transformée étant réalisée par blocs ou globalement.

19. Procédé selon une des revendications précédentes, caractérisé en ce qu'un nombre caractéristique est le contraste d'au moins une zone d'une image appartenant au flux de données.

20. Procédé selon une des revendications précédentes, caractérisé en ce qu'un nombre caractéristique est l'activité spatiale ou temporelle d'au moins une zone d'une image.

21. Procédé selon une des revendications précédentes, caractérisé en ce qu'un nombre caractéristique est la luminosité moyenne d'au moins une zone d'une image.

22. Procédé selon une des revendications précédentes, caractérisé en ce que lesdits points D_1 et D_2 sont générés à partir d'au moins deux paramètres caractéristiques.

23 Procédé selon la revendication 22, caractérisé en ce que lesdits paramètres caractéristiques sont des paramètres audio.

24. Procédé selon la revendication 22, caractérisé en ce que lesdits paramètres caractéristiques sont des paramètres vidéo.

25 Procédé selon une des revendications précédentes, caractérisé en ce que le flux de données comporte des données vidéo et des données audio et en ce qu'il met en œuvre une première synchronisation vidéo à partir de points D'_1 et D'_2 associées à au moins un paramètre caractéristique vidéo correspondant audit flux vidéo, et une deuxième synchronisation audio à partir de points D''_1 et D''_2 associés à au moins un paramètre caractéristique audio correspondant audit flux audio.

26. Procédé selon la revendication 25, caractérisé en ce qu'il présente une étape de détermination du décalage de synchronisation entre le flux vidéo et le flux audio par différence entre lesdits décalages obtenus pour le flux vidéo et pour le flux audio.

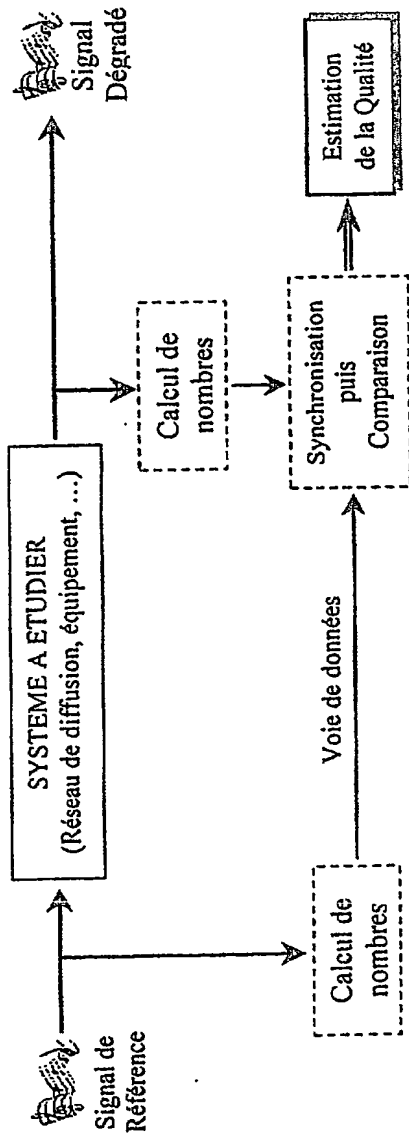


FIG. 1

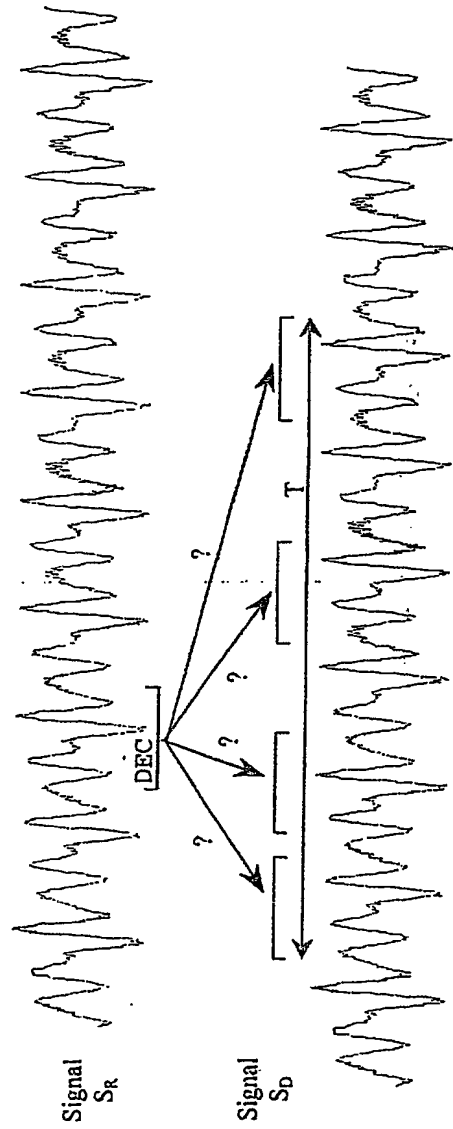


FIG. 2

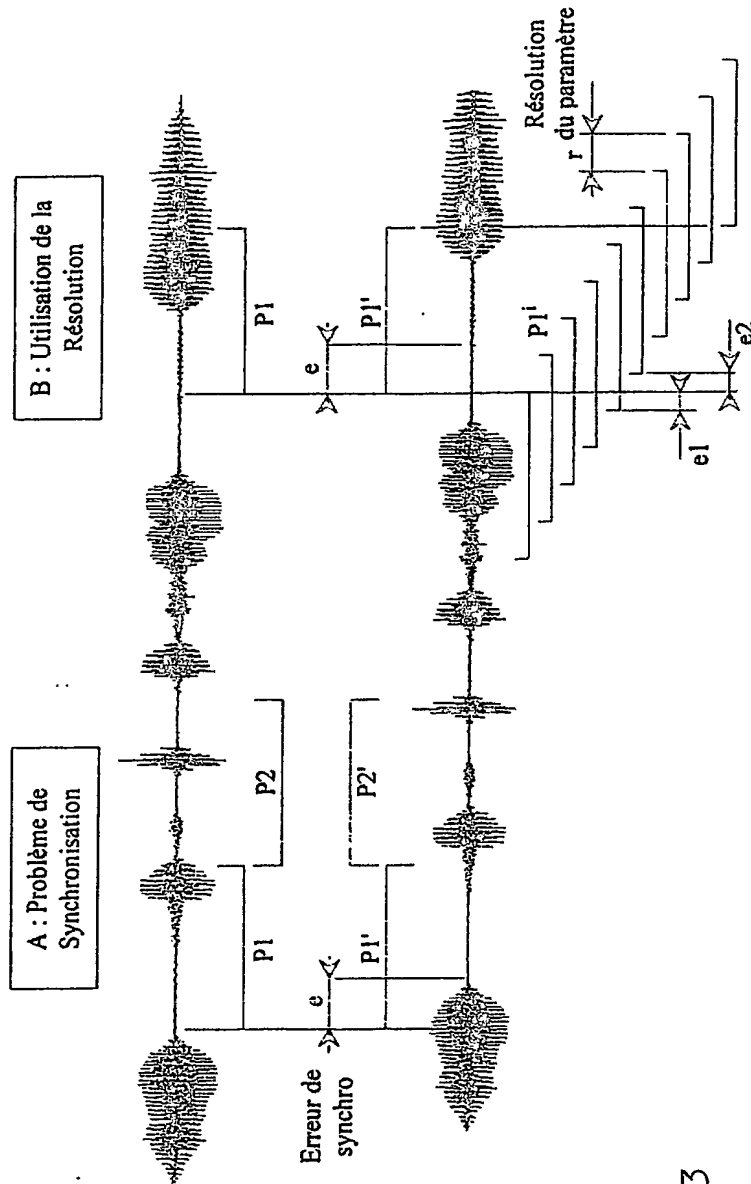


FIG. 3

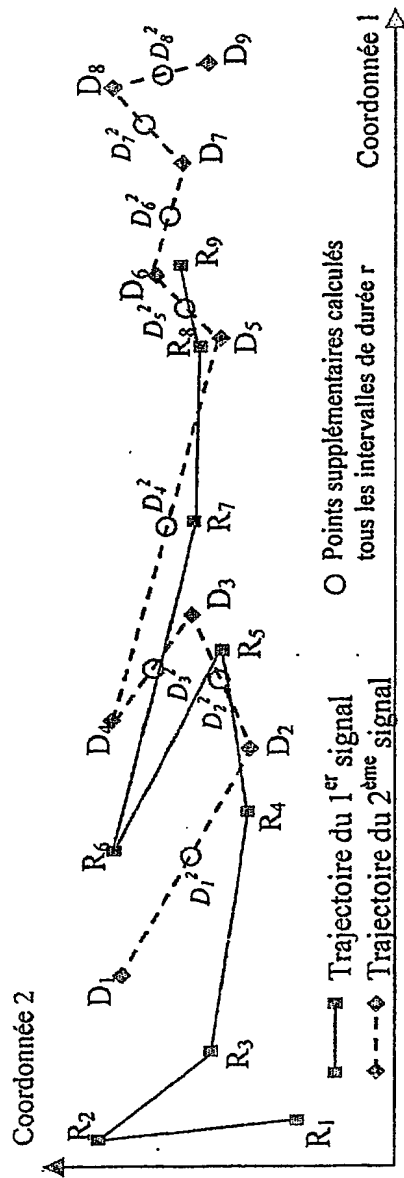


FIG. 4

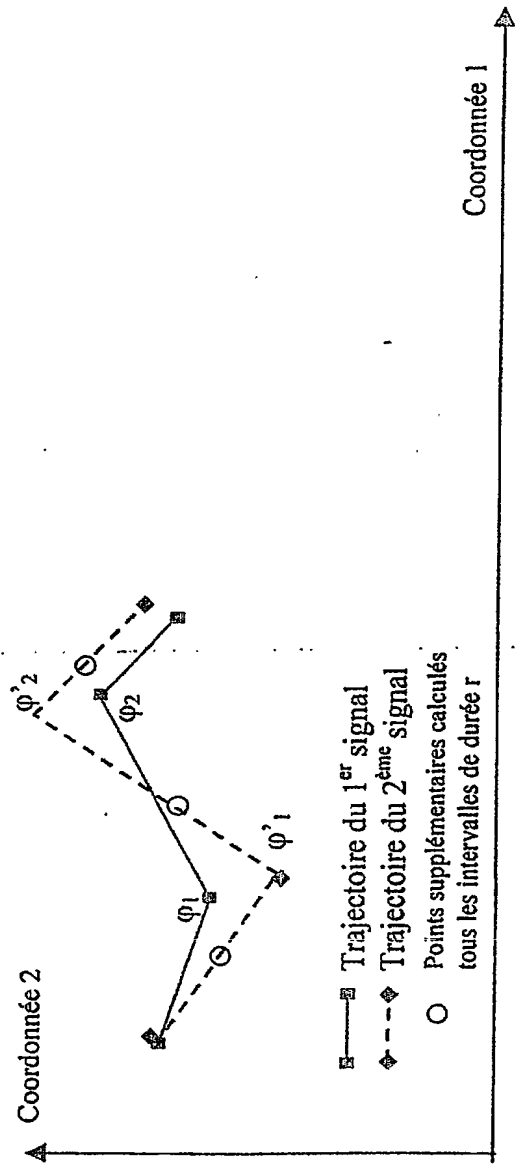


FIG. 5

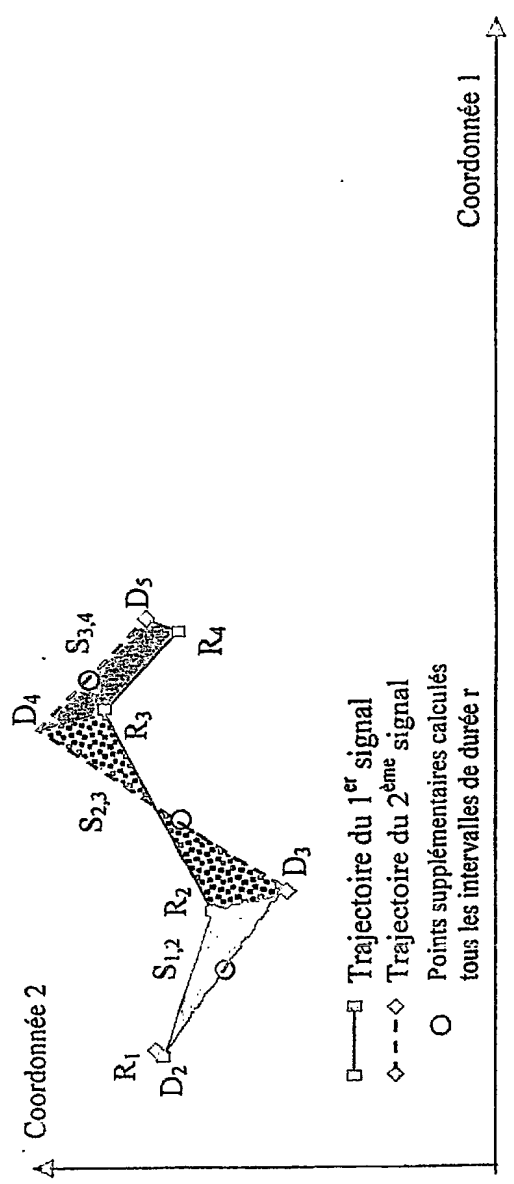


FIG. 6

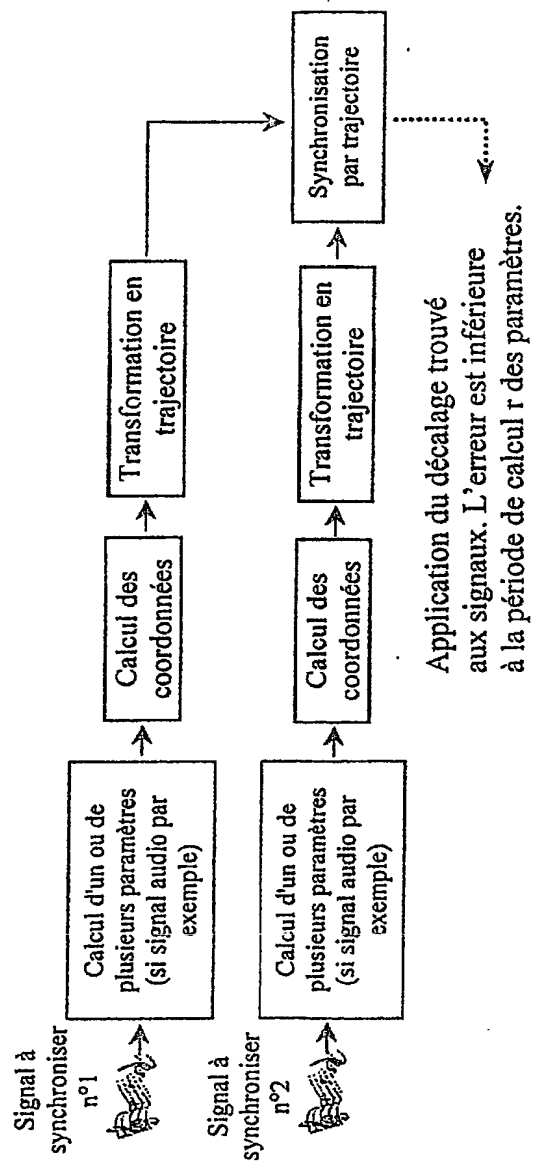


FIG. 7

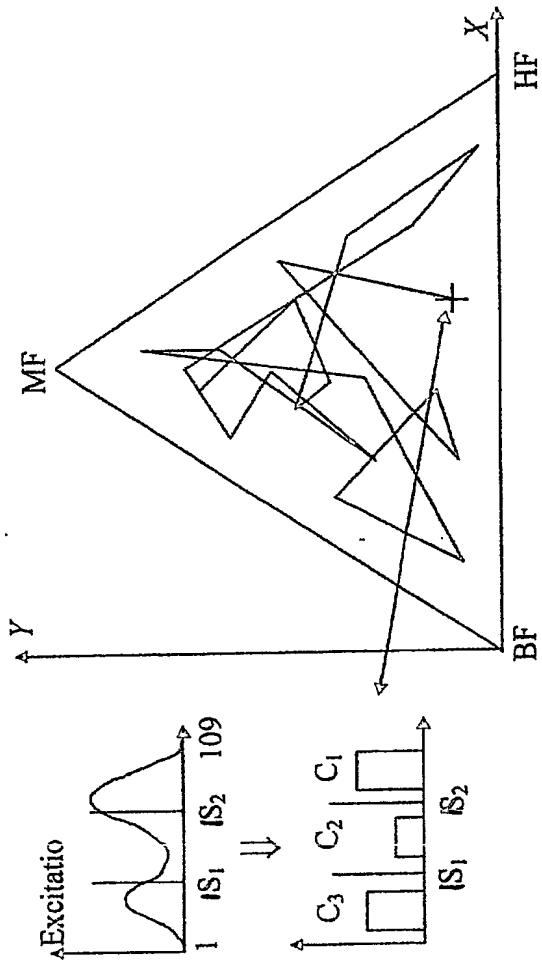


FIG. 8

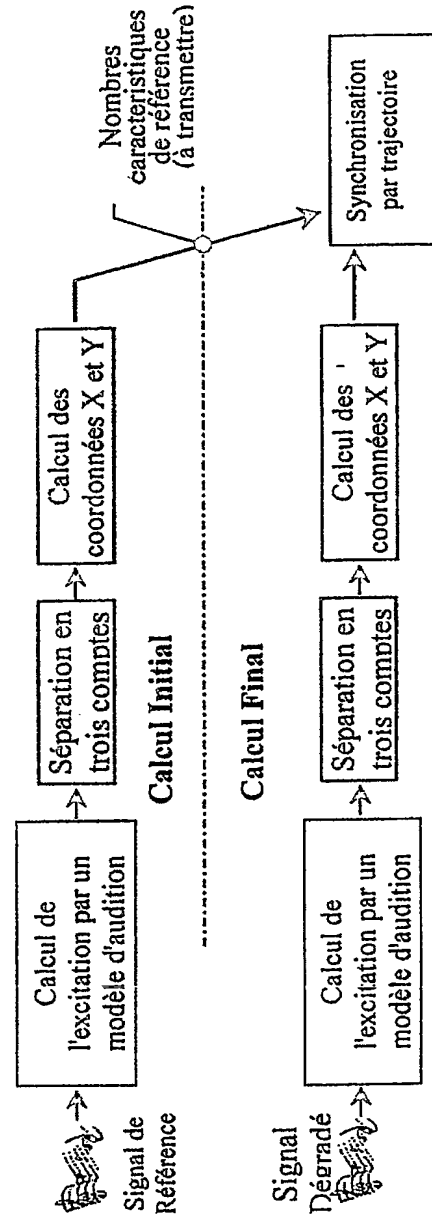
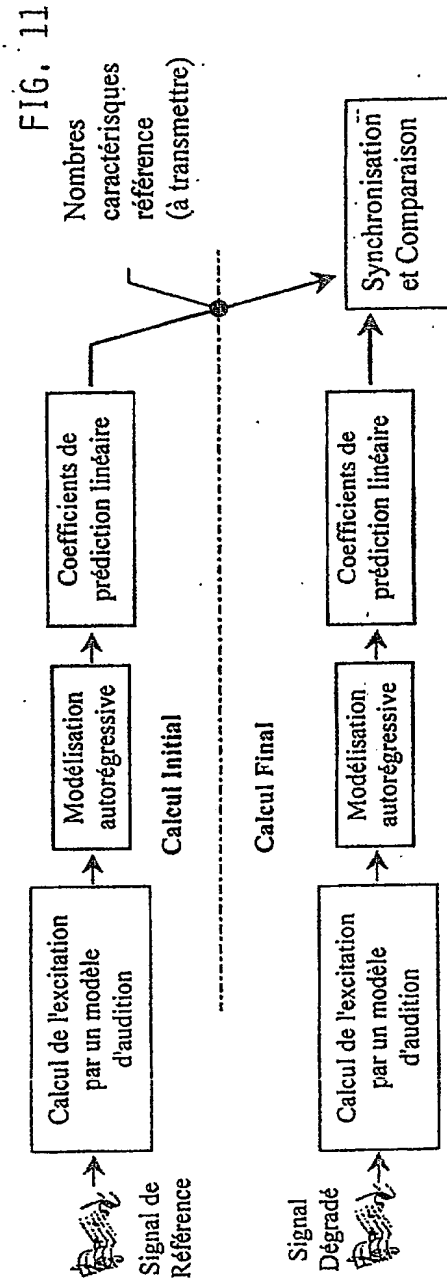
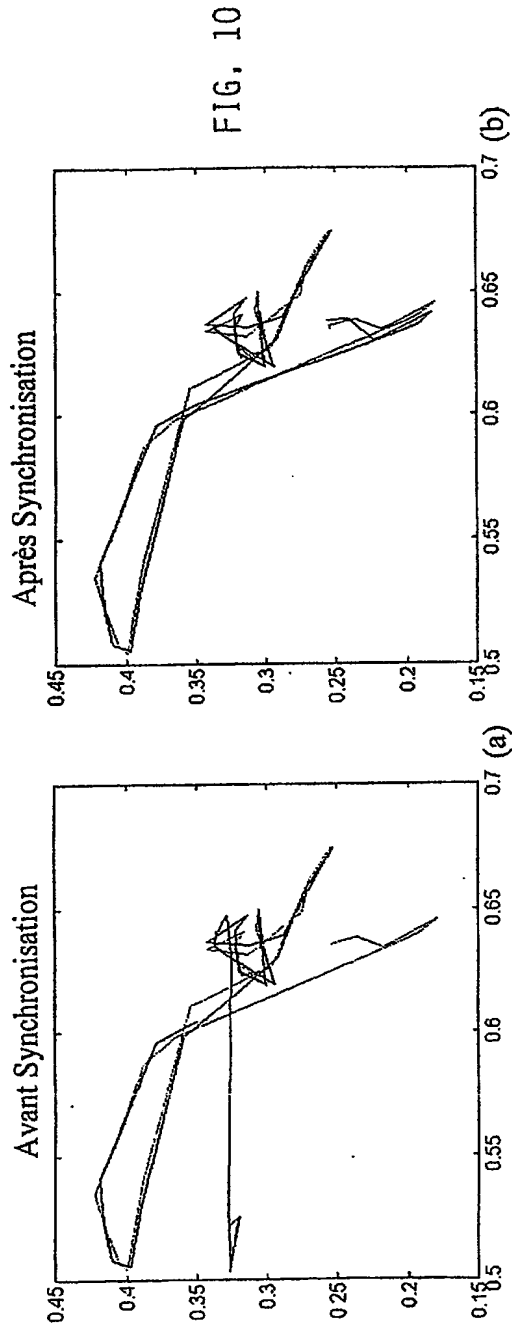


FIG. 9





BREVET D'INVENTION
CERTIFICAT D'UTILITÉ
Code de la propriété intellectuelle - Livre VI


N° 11 235 02


DÉPARTEMENT DES BREVETS

26 bis, rue de Saint Pétersbourg
75000 Paris Cedex 08
Téléphone : 01 53 04 53 04 Télécopie : 01 42 93 59 30

DÉSIGNATION D'INVENTEUR(S) Page N° 1. / 1.
(Si le demandeur n'est pas l'inventeur ou l'unique inventeur)

Cet imprimé est à remplir lisiblement à l'encre noire

DB 113 W / 260899

Vos références pour ce dossier (facultatif)		PJmnF1198/13 FR	
N° D'ENREGISTREMENT NATIONAL		0200872	
TITRE DE L'INVENTION (200 caractères ou espaces maximum)			
PROCÉDE DE SYNCHRONISATION DE DEUX FLUX DE DONNÉES NUMÉRIQUES DE MEME CONTENU.			
LE(S) DEMANDEUR(S) : TELEDIFFUSION DE FRANCE 10, rue d'Oradour-sur-Glanc 75015 PARIS FRANCE			
DESIGNE(NT) EN TANT QU'INVENTEUR(S) : (Indiquez en haut à droite «Page N° 1/1» S'il y a plus de trois inventeurs, utilisez un formulaire identique et numérotez chaque page en indiquant le nombre total de pages).			
Nom		JOLY	
Prénoms		Alexandre	
Adresse	Rue	61, rue de Tivoli	
	Code postal et ville	57070	METZ (France)
Société d'appartenance (facultatif)			
Nom			
Prénoms			
Adresse	Rue		
	Code postal et ville		
Société d'appartenance (facultatif)			
Nom			
Prénoms			
Adresse	Rue		
	Code postal et ville		
Société d'appartenance (facultatif)			
DATE ET SIGNATURE(S) DU (DES) DEMANDEUR(S) OU DU MANDATAIRE (Nom et qualité du signataire)		Paris, le 7 Mai 2002  JACQUARD Philippe - Mandataire n° 92-4024	

La loi n°78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés s'applique aux réponses faites à ce formulaire. Elle garantit un droit d'accès et de rectification pour les données vous concernant auprès de l'INPI.

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☒ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.